



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/024,955

DATE: 04/16/2002

TIME: 16:20:42

Input Set : N:\Crf3\RULE60\10024955.raw

Output Set: N:\CRF3\04162002\J024955.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Wayne R. Thomas and Kaw-Yan Chua

7 (ii) TITLE OF INVENTION: Allergenic Proteins and Peptides From  
8 House Dust Mite and Uses Therefor

10 (iii) NUMBER OF SEQUENCES: 15

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: LAHIVE & COCKFIELD, LLP

14 (B) STREET: 28 State Street

15 (C) CITY: Boston

16 (D) STATE: Massachusetts

17 (E) COUNTRY: USA

18 (F) ZIP: 02109

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/10/024,955

C--> 28 (B) FILING DATE: 19-Dec-2001

29 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: US/08/553,336A

34 (B) FILING DATE: 10-JUN-1996

36 (A) APPLICATION NUMBER: US 08/081,540

37 (B) FILING DATE: 22-JUNE-1993

39 (viii) ATTORNEY/AGENT INFORMATION:

40 (A) NAME: Jane E. Remillard

41 (B) REGISTRATION NUMBER: 38,872

42 (C) REFERENCE/DOCKET NUMBER: IMI-032CP2

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: (617)227-7400

46 (B) TELEFAX: (617)742-4214

49 (2) INFORMATION FOR SEQ ID NO: 1:

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 812 base pairs

53 (B) TYPE: nucleic acid

54 (C) STRANDEDNESS: single

55 (D) TOPOLOGY: linear

57 (ii) MOLECULE TYPE: cDNA

60 (ix) FEATURE:

61 (A) NAME/KEY: CDS

ENTERED

## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/10/024,955

TIME: 16:20:42

Input Set : N:\Crif3\RULE60\10024955.raw

Output Set: N:\CRF3\04162002\J024955.raw

```

62      (B) LOCATION: 68..712
64      (ix) FEATURE:
65          (A) NAME/KEY: mat_peptide
66          (B) LOCATION: 119..712
69      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
71 TTTTTTTTTT TTTGGTTAT TCCCATTTTT TTCATATCGT AAAAATCCAA ATTCACITTTT      60
73 TTACCAA ATG ATG AAA TTA TTA TTG ATT GCT GCC GCA GCT TTT GTT GCC      109
74      Met Met Lys Leu Leu Leu Ile Ala Ala Ala Ala Phe Val Ala
75      -17      -15      -10      -5
77 GTT TCG GCT GAT CCA ATT CAC TAT GAT AAA ATC ACC GAA GAA ATT AAC      157
78 Val Ser Ala Asp Pro Ile His Tyr Asp Lys Ile Thr Glu Glu Ile Asn
79      1      5      10
81 AAA GCT GTT GAT GAA GCC GTC GCT GCA ATT GAA AAA TCC GAA ACA TTC      205
82 Lys Ala Val Asp Glu Ala Val Ala Ala Ile Glu Lys Ser Glu Thr Phe
83      15      20      25
85 GAT CCA ATG AAG GTA CCC GAT CAT TCT GAT AAA TTC GAA CGA CAT ATT      253
86 Asp Pro Met Lys Val Pro Asp His Ser Asp Lys Phe Glu Arg His Ile
87      30      35      40      45
89 GGT ATC ATC GAT TTA AAA GGT GAA TTA GAC ATG CGA AAC ATT CAA GTT      301
90 Gly Ile Ile Asp Leu Lys Gly Glu Leu Asp Met Arg Asn Ile Gln Val
91      50      55      60
93 CGA GGA TTA AAA CAA ATG AAA CGT GTA GGT GAT GCT AAT GTG AAA AGT      349
94 Arg Gly Leu Lys Gln Met Lys Arg Val Gly Asp Ala Asn Val Lys Ser
95      65      70      75
97 GAA GAT GGT GTT GTC AAA GCT CAT TTG TTG GTC GGT GTT CAT GAT GAC      397
98 Glu Asp Gly Val Val Lys Ala His Leu Leu Val Gly Val His Asp Asp
99      80      85      90
101 GTT GTT TCA ATG GAA TAT GAT TTA GCA TAC AAA TTG GGT GAT CTT CAT      445
102 Val Val Ser Met Glu Tyr Asp Leu Ala Tyr Lys Leu Gly Asp Leu His
103      95      100      105
105 CCA AAC ACT CAT GTC ATT TCG GAT ATT CAG GAT TTT GTT GTC GAA TTA      493
106 Pro Asn Thr His Val Ile Ser Asp Ile Gln Asp Phe Val Val Glu Leu
107      110      115      120      125
109 TCG CTC GAA GTT AGC GAA GAA GGT AAT ATG ACA TTG ACA TCG TTC GAA      541
110 Ser Leu Glu Val Ser Glu Glu Gly Asn Met Thr Leu Thr Ser Phe Glu
111      130      135      140
113 GTA CGT CAA TTT GCC AAT GTT GTC AAT CAT ATT GGT GGT CTT TCA ATT      589
114 Val Arg Gln Phe Ala Asn Val Val Asn His Ile Gly Gly Leu Ser Ile
115      145      150      155
117 TTG GAT CCA ATT TTC GCT GTC TTA TCC GAT GTT TTG ACC GCT ATT TTC      637
118 Leu Asp Pro Ile Phe Ala Val Leu Ser Asp Val Leu Thr Ala Ile Phe
119      160      165      170
121 CAG GAT ACC GTA CGT GCA GAA ATG ACC AAA GTA TTG GCA CCA GCA TTC      685
122 Gln Asp Thr Val Arg Ala Glu Met Thr Lys Val Leu Ala Pro Ala Phe
123      175      180      185
125 AAA AAA GAA TTG GAA CGA AAC AAC CAA TAGACTTACA CACAACATAA      732
126 Lys Lys Glu Leu Glu Arg Asn Asn Gln
127      190      195
129 CACTGTTATT TTTACTGTTG ATAATCAAAT GAAATAAATT TTTTATCAT TTTGTTTAAA      792

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/024,955

DATE: 04/16/2002

TIME: 16:20:42

Input Set : N:\Crf3\RULE60\10024955.raw

Output Set: N:\CRF3\04162002\J024955.raw

131 AAAAAAAAAA AAAAAAAAAA 812

134 (2) INFORMATION FOR SEQ ID NO: 2:

136 (i) SEQUENCE CHARACTERISTICS:

137 (A) LENGTH: 215 amino acids

138 (B) TYPE: amino acid

139 (D) TOPOLOGY: linear

141 (ii) MOLECULE TYPE: protein

143 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

145 Met Met Lys Leu Leu Leu Ile Ala Ala Ala Ala Phe Val Ala Val Ser  
146 -17 -15 -10 -5  
148 Ala Asp Pro Ile His Tyr Asp Lys Ile Thr Glu Glu Ile Asn Lys Ala  
149 1 5 10 15  
151 Val Asp Glu Ala Val Ala Ala Ile Glu Lys Ser Glu Thr Phe Asp Pro  
152 20 25 30  
154 Met Lys Val Pro Asp His Ser Asp Lys Phe Glu Arg His Ile Gly Ile  
155 35 40 45  
157 Ile Asp Leu Lys Gly Glu Leu Asp Met Arg Asn Ile Gln Val Arg Gly  
158 50 55 60  
160 Leu Lys Gln Met Lys Arg Val Gly Asp Ala Asn Val Lys Ser Glu Asp  
161 65 70 75  
163 Gly Val Val Lys Ala His Leu Leu Val Gly Val His Asp Asp Val Val  
164 80 85 90 95  
166 Ser Met Glu Tyr Asp Leu Ala Tyr Lys Leu Gly Asp Leu His Pro Asn  
167 100 105 110  
169 Thr His Val Ile Ser Asp Ile Gln Asp Phe Val Val Glu Leu Ser Leu  
170 115 120 125  
172 Glu Val Ser Glu Glu Gly Asn Met Thr Leu Thr Ser Phe Glu Val Arg  
173 130 135 140  
175 Gln Phe Ala Asn Val Val Asn His Ile Gly Gly Leu Ser Ile Leu Asp  
176 145 150 155  
178 Pro Ile Phe Ala Val Leu Ser Asp Val Leu Thr Ala Ile Phe Gln Asp  
179 160 165 170 175  
181 Thr Val Arg Ala Glu Met Thr Lys Val Leu Ala Pro Ala Phe Lys Lys  
182 180 185 190  
184 Glu Leu Glu Arg Asn Asn Gln  
185 195

187 (2) INFORMATION FOR SEQ ID NO: 3:

189 (i) SEQUENCE CHARACTERISTICS:

190 (A) LENGTH: 18 base pairs

191 (B) TYPE: nucleic acid

192 (C) STRANDEDNESS: single

193 (D) TOPOLOGY: linear

195 (ii) MOLECULE TYPE: cDNA

199 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

201 GATCCAATTC ACTATGAT 18

203 (2) INFORMATION FOR SEQ ID NO: 4:

205 (i) SEQUENCE CHARACTERISTICS:

206 (A) LENGTH: 17 base pairs

207 (B) TYPE: nucleic acid

## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/10/024,955

TIME: 16:20:42

Input Set : N:\Crif3\RULE60\10024955.raw

Output Set: N:\CRF3\04162002\J024955.raw

```

208      (C) STRANDEDNESS: single
209      (D) TOPOLOGY: linear
211      (ii) MOLECULE TYPE: cDNA
215      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
217 GGTGAATTAG ACATGCG                                     17
219 (2) INFORMATION FOR SEQ ID NO: 5:
221      (i) SEQUENCE CHARACTERISTICS:
222          (A) LENGTH: 24 base pairs
223          (B) TYPE: nucleic acid
224          (C) STRANDEDNESS: single
225          (D) TOPOLOGY: linear
227      (ii) MOLECULE TYPE: cDNA
231      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
233 TCAATTTTGG ATCCAATTTT CGCT                               24
235 (2) INFORMATION FOR SEQ ID NO: 6:
237      (i) SEQUENCE CHARACTERISTICS:
238          (A) LENGTH: 761 base pairs
239          (B) TYPE: nucleic acid
240          (C) STRANDEDNESS: single
241          (D) TOPOLOGY: linear
243      (ii) MOLECULE TYPE: cDNA
246      (ix) FEATURE:
247          (A) NAME/KEY: CDS
248          (B) LOCATION: 43..681
251      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
253 GATCTTATAT CAATAACAAT CCAAAAAAAC ATATCTTACA AA ATG ATG AAA TTT       54
254                                     Met Met Lys Phe
255                                     1
257 TTG TTG ATT GCT GCC GTG GCA TTT GTC GCC GTT TCG GCT GAT CCA ATT       102
258 Leu Leu Ile Ala Ala Val Ala Phe Val Ala Val Ser Ala Asp Pro Ile
259   5          10          15          20
261 CAC TAT GAT AAA ATC ACC GAA GAA ATC AAC AAA GCT ATT GAT GAT GCC       150
262 His Tyr Asp Lys Ile Thr Glu Glu Ile Asn Lys Ala Ile Asp Asp Ala
263          25          30          35
265 ATT GCT GCT ATT GAA CAA TCC GAA ACA ATA GAT CCA ATG AAA GTA CCT       198
266 Ile Ala Ala Ile Glu Gln Ser Glu Thr Ile Asp Pro Met Lys Val Pro
267          40          45          50
269 GAT CAT GCC GAT AAA TTC GAA CGT CAT GTT GGT ATT GTG GAT TTC AAA       246
270 Asp His Ala Asp Lys Phe Glu Arg His Val Gly Ile Val Asp Phe Lys
271          55          60          65
273 GGT GAA TTA GCC ATG CGA AAC ATT GAG GCT CGA GGA TTG AAA CAA ATG       294
274 Gly Glu Leu Ala Met Arg Asn Ile Glu Ala Arg Gly Leu Lys Gln Met
275          70          75          80
277 AAA CGT CAA GGT GAT GCT AAT GTC AAA GGT GAA GAG GGT ATT GTT AAA       342
278 Lys Arg Gln Gly Asp Ala Asn Val Lys Gly Glu Glu Gly Ile Val Lys
279          85          90          95          100
281 GCT CAT TTG TTG ATC GGT GTT CAC GAT GAT ATC GTC TCG ATG GAA TAT       390
282 Ala His Leu Leu Ile Gly Val His Asp Asp Ile Val Ser Met Glu Tyr
283          105          110          115

```

## RAW SEQUENCE LISTING

DATE: 04/16/2002

PATENT APPLICATION: US/10/024,955

TIME: 16:20:42

Input Set : N:\Crf3\RULE60\10024955.raw

Output Set: N:\CRF3\04162002\J024955.raw

```

285 GAT TTA GCA TAC AAA TTG GGT GAT CTT CAT CCA ACC ACT CAT GTC ATT      438
286 Asp Leu Ala Tyr Lys Leu Gly Asp Leu His Pro Thr Thr His Val Ile
287           120           125           130
289 TCG GAT ATT CAA GAT TTT GTT GTT GCC TTG TCC CTT GAA ATT TCT GAT      486
290 Ser Asp Ile Gln Asp Phe Val Val Ala Leu Ser Leu Glu Ile Ser Asp
291           135           140           145
293 GAA GGT AAC ATA ACA ATG ACA TCT TTT GAA GTA CGA CAA TTC GCT AAT      534
294 Glu Gly Asn Ile Thr Met Thr Ser Phe Glu Val Arg Gln Phe Ala Asn
295           150           155           160
297 GTT GTC AAC CAT ATT GGT GGT CTT TCA ATC TTG GAT CCA ATT TTT GGC      582
298 Val Val Asn His Ile Gly Gly Leu Ser Ile Leu Asp Pro Ile Phe Gly
299 165           170           175           180
301 GTT TTA TCT GAT GTA TTG ACC GCT ATT TTC CAA GAC ACC GTA CGT AAG      630
302 Val Leu Ser Asp Val Leu Thr Ala Ile Phe Gln Asp Thr Val Arg Lys
303           185           190           195
305 GAA ATG ACC AAA GTA TTG GCA CCA GCA TTT AAA CGT GAA TTG GAA AAA      678
306 Glu Met Thr Lys Val Leu Ala Pro Ala Phe Lys Arg Glu Leu Glu Lys
307           200           205           210
309 AAT TAACCAATAG ACATCATTTT TCCAACGTGA CAATCTCTAT TTCACTGACA      731
310 Asn
313 ATAAATAAAA ATTTTATTTT TTATTTCTCC      761
316 (2) INFORMATION FOR SEQ ID NO: 7:
318     (i) SEQUENCE CHARACTERISTICS:
319         (A) LENGTH: 213 amino acids
320         (B) TYPE: amino acid
321         (D) TOPOLOGY: linear
323     (ii) MOLECULE TYPE: protein
325     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
327 Met Met Lys Phe Leu Leu Ile Ala Ala Val Ala Phe Val Ala Val Ser
328   1           5           10           15
330 Ala Asp Pro Ile His Tyr Asp Lys Ile Thr Glu Glu Ile Asn Lys Ala
331           20           25           30
333 Ile Asp Asp Ala Ile Ala Ala Ile Glu Gln Ser Glu Thr Ile Asp Pro
334           35           40           45
336 Met Lys Val Pro Asp His Ala Asp Lys Phe Glu Arg His Val Gly Ile
337           50           55           60
339 Val Asp Phe Lys Gly Glu Leu Ala Met Arg Asn Ile Glu Ala Arg Gly
340 65           70           75           80
342 Leu Lys Gln Met Lys Arg Gln Gly Asp Ala Asn Val Lys Gly Glu Glu
343           85           90           95
345 Gly Ile Val Lys Ala His Leu Leu Ile Gly Val His Asp Asp Ile Val
346           100          105          110
348 Ser Met Glu Tyr Asp Leu Ala Tyr Lys Leu Gly Asp Leu His Pro Thr
349           115          120          125
351 Thr His Val Ile Ser Asp Ile Gln Asp Phe Val Val Ala Leu Ser Leu
352           130          135          140
354 Glu Ile Ser Asp Glu Gly Asn Ile Thr Met Thr Ser Phe Glu Val Arg
355 145          150          155          160
357 Gln Phe Ala Asn Val Val Asn His Ile Gly Gly Leu Ser Ile Leu Asp

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/024,955

DATE: 04/16/2002

TIME: 16:20:43

Input Set : N:\Crf3\RULE60\10024955.raw

Output Set: N:\CRF3\04162002\J024955.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]